

## ON A PROBLEM BY SHAPOZENKO ON JOHNSON GRAPHS

VÍCTOR DIEGO, ORIOL SERRA, AND LLUÍS VENA

**ABSTRACT.** The Johnson graph  $J(n, m)$  has the  $m$ -subsets of  $\{1, 2, \dots, n\}$  as vertices and two subsets are adjacent in the graph if they share  $m - 1$  elements. Shapozenko asked about the isoperimetric function  $\mu_{n,m}(k)$  of Johnson graphs, that is, the cardinality of the smallest boundary of sets with  $k$  vertices in  $J(n, m)$  for each  $1 \leq k \leq \binom{n}{m}$ . We give an upper bound for  $\mu_{n,m}(k)$  and show that, for each given  $k$  such that the solution to the Shadow Minimization Problem in the Boolean lattice is unique, and each sufficiently large  $n$ , the given upper bound is tight. We also show that the bound is tight for the small values of  $k \leq m + 1$  and for all values of  $k$  when  $m = 2$ .

**Keywords:** Johnson graph, Isoperimetric problem, Shift compression.

## 1. INTRODUCTION

Let  $G = (V, E)$  be a graph. Given a set  $X \subset V$  of vertices, we denote by

$$\partial X = \{y \in V \setminus X : d(X, y) = 1\}, \quad B(X) = \{y \in V : d(X, y) \leq 1\} = X \cup \partial X,$$

the *boundary* and the *ball* of  $X$  respectively, where  $d(X, y)$  denotes  $\min\{d(x, y) : x \in X\}$ .

We write  $\partial_G$  and  $B_G$  when the reference to  $G$  has to be made explicit. The *vertex-isoperimetric function* (we will call it simply *isoperimetric function*) of  $G$  is defined as

$$\mu_G(k) = \min\{|\partial X| : X \subset V, |X| = k\},$$

that is,  $\mu_G(k)$  is the size of the smallest boundary among sets of vertices with cardinality  $k$ .

The isoperimetric function is known only for a few classes of graphs. One of the seminal results is the exact determination of the isoperimetric function for the  $n$ -cube obtained by Harper [17] in 1966 (and by Hart with the edge-isoperimetric function at [19] in 1976.) Analogous results were obtained for cartesian products of chains by Bollobás and Leader [8] and Bezrukov [4], cartesian products of even

---

V. Diego acknowledges support from *Ministerio de Ciencia e Innovación*, Spain, and the *European Regional Development Fund* under a FPI grant in the project MTM2011-28800-C02-01.

O. Serra was supported by the Spanish Ministerio de Economía y Competitividad under project MTM2014-54745-P. .

L. Vena was supported by the Center of Excellence-Inst. for Theor. Comp. Sci., Prague, P202/12/G061, and by Project ERCCZ LL1201 *CORES*.

cycles by Karachanjan [20] and Riordan [25] (see also Bezrukov and Leck at [5]) and some other cartesian products by Bezrukov and Serra [6].

The Johnson graph  $J(n, m)$  has the  $m$ -subsets of  $[n] = \{1, 2, \dots, n\}$  as vertices and two  $m$ -subsets are adjacent in the graph whenever their symmetric difference has cardinality 2. It follows from the definition that, for  $m = 1$ , the Johnson graph  $J(n, 1)$  is the complete graph  $K_n$ . For  $m = 2$  the Johnson graph  $J(n, 2)$  is the line graph of the complete graph on  $n$  vertices, also known as the triangular graph  $T(n)$ . Thus, for instance,  $J(5, 2)$  is the complement of the Petersen graph, displayed in Figure 1. Also,  $J(n, 2)$  is the complement of the Kneser graph  $K(n, 2)$ , the graph which has the 2-subsets of  $[n]$  as vertices and two pairs are adjacent whenever they are disjoint.

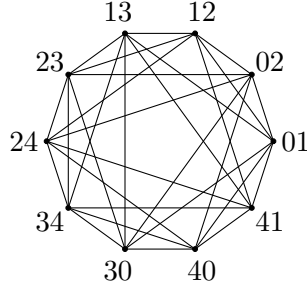


FIGURE 1. The Johnson graph  $J(5, 2)$ .

Johnson graphs arise from the association schemes named after Johnson who introduced them, see e.g. [10]. The Johnson graphs are one of the important classes of distance-transitive graphs; see e.g. Brouwer, Cohen, Neumaier [1, Chapter 9] or Godsil [16, Chapter 11].

Given a family  $S$  of  $m$ -sets of an  $n$ -set, its *lower shadow*  $\Delta(S)$  is the family of  $(m - 1)$ -sets which are contained in some  $m$ -set in  $S$ . The *upper shadow*  $\nabla(S)$  of  $S$  is the family of  $(m + 1)$ -sets which contain some  $m$ -set in  $S$ . The *ball* of  $S$  in the Johnson graph  $J(n, m)$  can be written as

$$B(S) = \nabla(\Delta(S)) = \Delta(\nabla(S)). \quad (1)$$

These equalities establish a connection between the isoperimetric problem in the Johnson graph with the Shadow Minimization Problem (SMP) in the Boolean lattice, which consists in finding, for a given  $k$ , the smallest cardinality of  $\Delta(S)$  among all families  $S$  of  $m$ -sets with cardinality  $k$ . The latter problem is solved by the well-known Kruskal–Katona theorem [22, 21], which establishes that the initial segments in the colex order provide a family of extremal sets for the SMP.

Recall that the *colex order* in the set of  $m$ -subsets of  $[n]$  is defined as  $X \leq Y$  if and only if  $\max((X \setminus Y) \cup (Y \setminus X)) \in Y$  (we follow here the terminology from Bollobás

[7, Section 5]; we also use  $\binom{[n]}{m}$  to denote the family of  $m$ -subsets of an  $n$ -set, and  $[k, l] = \{k, k+1, \dots, l\}$  for integers  $k < l$ .) The computation of the boundary of initial segments in the colex order (the family of the first  $m$ -subsets in this order) provides the following upper bound for the isoperimetric function of Johnson graphs:

**Proposition 1.1.** *Let  $\mu_{n,m} : [N] \rightarrow \mathbb{N}$  denote the isoperimetric function of the Johnson graph  $J(n, m)$ , where  $N = \binom{n}{m}$ . Let*

$$k = \binom{k_0}{m} + \binom{k_1}{m-1} + \dots + \binom{k_r}{m-r}, \quad k_0 > \dots > k_r \geq m-r > 0,$$

*be the  $m$ -binomial representation of  $k$ . Then*

$$\mu_{n,m}(k) \leq f(k, n, m), \quad (2)$$

*where*

$$f(k, n, m) = \binom{k_0}{m-1}(n-k_0) + \sum_{i=1}^r \left( \binom{k_i}{m-i-1}(n-k_0-1) - \binom{k_i}{m-i} \right). \quad (3)$$

*Proof.* The initial segment  $I$  of length  $k$  in the colex order is the disjoint union

$$I = I_0 \cup \dots \cup I_r,$$

where  $I_0$  consists of all  $m$ -sets in  $\binom{[k_0]}{m}$  and, for  $j > 0$ ,  $I_j$  consists of all sets containing  $\{k_{j-1}+1, \dots, k_0+1\}$  and  $m-j$  elements in  $[k_j]$ . The right hand side of (3) is the cardinality of  $\partial I$  as can be shown by induction on  $r$ . If  $r = 0$  then  $\partial I$  consists of the  $\binom{k_0}{m-1}(n-k_0)$  sets obtained by replacing one element in  $[k_0]$  by one element in  $[k_0+1, n]$  from a set in  $I$ . Suppose that  $r > 0$  and write  $I = I' \cup I_r$  as the disjoint union of  $I' = I_0 \cup \dots \cup I_{r-1}$  and  $I_r$ . We have  $\partial I = (\partial I' \setminus I_r) \cup (\partial I_r \setminus B(I'))$ , the union being disjoint. Since  $I_r \subset \partial I'$  we have,

$$|\partial I' \setminus I_r| = |\partial I'| - |I_r| = |\partial I'| - \binom{k_r}{m-r},$$

while

$$|\partial I_r \setminus B(I')| = \binom{k_r}{m-r-1}(n-k_0-1),$$

since the only sets in  $\partial I_r \setminus B(I')$  are those obtained from a set in  $I_r$  by replacing one element in  $[k_r]$  by one element in  $[k_0+2, n]$ .  $\square$

The family of initial segments in the colex order does not provide in general a solution to the isoperimetric problem in  $J(n, m)$ . A simple example is as follows.

**Example 1.1.** *Take  $n = 3(m+1)/2$ . The ball  $B(\{\mathbf{x}\})$  of radius one in  $J(3(m+1)/2, m)$  has cardinality*

$$|B_1| = 1 + m(n-m) = \frac{(m+2)(m+1)}{2} = \binom{m+2}{m},$$

and its boundary has cardinality

$$|\partial B_1| = \binom{m}{2} \binom{n-m}{2} = \frac{m(m-1)(m+3)(m+1)}{16}.$$

On the other hand, according to (3) and the  $m$ -binomial decomposition of  $|B_1|$ , the initial segment  $I$  of length  $|B_1|$  has cardinality

$$\begin{aligned} |\partial I| &= \binom{m+2}{m-1} \frac{m-1}{2} = \frac{(m+2)(m+1)m(m-1)}{12} \\ &= |\partial B_1| + \frac{(m+1)m(m-1)^2}{48}, \end{aligned}$$

which shows that the unit ball can have, as a function of  $m$ , an arbitrarily smaller boundary than the initial segment in the colex order.

In his monograph on discrete isoperimetric problems Leader [23] mentions the isoperimetric problem for Johnson graphs as one of the intriguing open problems in the area. Later on, in his extensive monograph on isoperimetric problems, Harper [18] attributes the problem to Shapozenko, and recalls that it is still open. Recently, Christofides, Ellis and Keevash [11] have obtained a lower bound for the isoperimetric function of Johnson graphs which is asymptotically tight for sets with cardinality  $\frac{1}{2} \binom{n}{m}$ . The Johnson graphs  $J(n, 2)$  provide a counterexample to a conjecture of Brouwer on the 2-restricted connectivity of strongly regular graphs, see Cioabă, Kim and Koolen [27] and Cioabă, Koolen and Li [26], where the connectivity of the more general class of strongly regular graphs and distance-regular graphs is studied. It is also worth mentioning that the edge version of the isoperimetric problem, where the minimization is for the number of edges leaving a set of given cardinality, has also been studied, see e.g. Ahlswede and Katona [2] or Bey [3]. We will only deal with the vertex isoperimetric problem in this paper and refer to it simply as the isoperimetric problem.

Our main purpose in this paper is to show that the initial segments in the colex order still provide a solution to the isoperimetric problem in  $J(n, m)$  for many small values of  $k$ , thus providing the exact value of the isoperimetric function in these cases.

We call a set  $S$  of vertices of  $J(n, m)$  *optimal* if  $|\partial(S)| = \mu_{n,m}(|S|)$ . Our first result shows that initial segments in the colex order are optimal sets in  $J(n, 2)$ .

**Theorem 1.2.** *For each  $n \geq 3$  and each  $1 \leq k \leq \binom{n}{2}$  we have*

$$\mu_{n,2}(k) = f(k, n, 2).$$

*In particular, the initial segments in the colex order are optimal sets of  $J(n, 2)$  for each  $n \geq 3$ .*

The following theorem allows one to show that the inequality (2) is also tight in  $J(n, m)$  for very small sets.

**Theorem 1.3.** *For  $k < m - 1$  and  $n \geq 2(m - 1)$  the initial segment of length  $k$  of the colex order in  $J(n, m)$  is an optimal set.*

Our last result, Theorem 1.4, extends Theorem 1.3 in an asymptotic way, by showing that the inequality (2) is tight for a large number of small cardinalities and gives a lower bound for all small cardinalities.

**Theorem 1.4.** *Let  $k, m$  be positive integers and let*

$$k = \binom{k_0}{m} + \binom{k_1}{m-1} + \cdots + \binom{k_r}{m-r}, \quad k_0 > \cdots > k_r \geq m-r > 0,$$

*be the  $m$ -binomial representation of  $k$ .*

*There is  $n(k, m)$  such that, for all  $n \geq n(k, m)$ , the following holds.*

(i) *If  $r < m - 1$  then*

$$\mu_{n,m}(k) = f(k, n, m),$$

*and the initial segment in the colex order with length  $k$  is the only (up to automorphisms) optimal set with cardinality  $k$  of the Johnson graph  $J(n, m)$ .*

(ii) *If  $r = m - 1$  then*

$$\mu_{n,m}(k) \leq f(k - k_r, n, m) + k_r.$$

The proof of Theorem 1.4 provides the estimation

$$n(k, m) \leq m + k + 1 - \mu_{m+k+1,m}(k) + f(k, m + k + 1, m)$$

for the value of  $n(k, m)$  above for which the statement of Theorem 1.4 holds. This upper bound for  $n(k, m)$  is not tight but we make no attempt to optimize its value in this paper.

Example 1.1 shows that the initial segment in the colex order of length  $\binom{m+2}{m}$  can fail to be an optimal set in  $J(n, m)$  if  $n = 3(m+1)/2$ . In the last section we describe another infinite family of examples for which the initial segment in the colex order fails again to be an optimal set in  $J(n, m)$  for every fixed  $m$  and all  $n$  large enough.

**Proposition 1.5.** *Let  $m$  be a positive integer. For each integer  $k$  of the form*

$$k = \binom{t}{m} + 3 \binom{t}{m-1}$$

*with  $t$  sufficiently large with respect to  $m$  there is a set  $S$  with cardinality  $k$  such that*

$$|B(S)| < f(k, n, m)$$

*for all  $n \geq t + 3$ .*

When  $n = t + 3$  the set  $S$  in Proposition 1.5 can be easily described as the ball in  $J(n, m)$  of  $\binom{[t]}{m}$ , the family of  $m$ -subsets of the first  $t$  symbols. Such sets are clear candidates to be optimal sets. The examples described in Proposition 1.5 are closely

related to the non-uniqueness of solutions to the Shadow Minimization Problem in the Boolean lattice (see Theorem 1.6 below.)

Standard compression techniques are used to prove the above results. These tools fall short to solve the isoperimetric problem of Johnson graphs in full mainly because, as pointed out in [11], for instance, optimal sets in Johnson graphs do not have the nested property (the ball of an optimal set is not optimal.) However these techniques are still useful to show that the colex order provides a sequence of extremal sets for small cardinalities.

The paper is organized as follows. Section 2 recalls the shifting techniques and compression of sets. The proofs of theorems 1.2, 1.3 and 1.4 are given in sections 3, 4 and 5 respectively. In the proof of Theorem 1.4 we use a result by Füredi and Griggs [14] which characterizes the cardinalities for which the Shadow Minimization Problem for the Boolean lattice has unique solution. The statement below is a rewriting of a combination of Proposition 2.3 and Theorem 2.6 in [14].

**Theorem 1.6** ([14]). *Let*

$$k = \binom{k_0}{m} + \binom{k_1}{m-1} + \cdots + \binom{k_r}{m-r}, \quad k_0 > \cdots > k_r \geq m-r > 0,$$

*be the  $m$ -binomial representation of  $k$ .*

*The initial segment in the colex order is the unique (up to automorphisms) solution to the Shadow Minimization Problem in the Boolean lattice if and only if  $r < m-1$ .*

Finally in Section 6 we prove Proposition 1.5. The result describes an infinite family of examples which show that the initial segments in colex order may fail to be optimal sets. The nature of this example shows that the isoperimetric problem in Johnson graphs still has many intriguing open questions to be solved.

## 2. SHIFTING TECHNIQUES

Shifting techniques are one of the key tools in the study of set systems. They were initially introduced in the original proof of the Erdős–Ko–Rado theorem [12] and have been particularly used in the solution by Frankl and Füredi [13] of the isoperimetric problem for hypercubes.

In what follows we identify subsets of  $[n]$  with their characteristic vectors  $\mathbf{x} = (x_1, \dots, x_n) \in \{0, 1\}^n$  where  $x_i = 1$  if  $i$  is in the corresponding set and  $x_i = 0$  otherwise. We denote the *support* of  $\mathbf{x}$  by

$$\overline{\mathbf{x}} = \{i : x_i = 1\}.$$

and the  $\ell_1$ -norm of  $\mathbf{x}$  by

$$|\mathbf{x}| = \sum_i x_i.$$

The support  $\overline{S}$  of  $S \subset \{0, 1\}^n$  is the union of the supports of its vectors. We often identify a set  $S \subset \{0, 1\}^n$  with the subset in  $\{0, 1\}^{n'}$ ,  $n' > n$ , obtained by adding zeros to the right in the coordinates of its vectors. Thus, the initial segment of length  $k$  is considered to be a subset of  $\{0, 1\}^n$  for each sufficiently large  $n$ .

The sum  $\mathbf{x} + \mathbf{y} = (x_1 + y_1 \pmod{2}, \dots, x_n + y_n \pmod{2})$ , of characteristic vectors is meant to be performed in the field  $\mathbb{F}_2^n$  and it corresponds to the symmetric difference of the corresponding sets. We also denote by  $\mathbf{e}_1, \dots, \mathbf{e}_n$  the unit vectors with 1 in the  $i$ -th coordinate and zero everywhere else.

With the above notation, the set of vertices of the Johnson graph  $J(n, m)$  are all vectors of  $\{0, 1\}^n$  with norm  $m$ , and the neighbors of  $\mathbf{x}$  in  $J(n, m)$  are the vectors

$$\mathbf{x} + \mathbf{e}_i + \mathbf{e}_j,$$

for each pair  $i, j$  such that  $x_i + x_j = 1$ .

We next recall the definition of the shifting transformation.

**Definition 2.1.** Let  $i, j \in [n]$ . For a set  $S \subset \{0, 1\}^n$  define

$$S_{ij} = \{\mathbf{x} \in S : x_i = 1 \text{ and } x_j = 0\},$$

and

$$T_{ij}(\mathbf{x}, S) = \begin{cases} \mathbf{x} + \mathbf{e}_i + \mathbf{e}_j, & \text{if } \mathbf{x} \in S_{ij} \text{ and } \mathbf{x} + \mathbf{e}_i + \mathbf{e}_j \notin S \\ \mathbf{x} & \text{otherwise} \end{cases}$$

The  $ij$ -shift of  $S$  is defined as

$$T_{ij}(S) = \{T_{ij}(\mathbf{x}, S) : \mathbf{x} \in S\}.$$

It follows from the definition that the shifting  $T_{ij}$  of a set preserves its cardinality and the norm of its elements. Moreover, it sends every vertex to a vertex at distance at most 1. The main property of the shifting transformation is that it does not increase the cardinality of the ball of a set. This property follows from the analogous ones for upper and lower shadows. We include a direct proof here for completeness.

**Lemma 2.1.** Let  $i, j \in [n]$  and write  $T = T_{ij}$ . For each set  $S$  of vertices in the Johnson graph  $J(n, m)$  we have

$$B(T(S)) \subseteq T(B(S)). \quad (4)$$

In particular,

$$|B(T(S))| \leq |T(B(S))| = |B(S)|. \quad (5)$$

*Proof.* We will show that,

$$\text{for each } \mathbf{y} \in T(S), \text{ we have } B(\mathbf{y}) \subseteq T(B(S)), \quad (6)$$

which is equivalent to (4). We observe that then (5) follows since  $|B(X)| = |X| + |\partial X|$  for every subset  $X$  and

$$|\partial(T(S))| = |B(T(S))| - |T(S)| \leq |T(B(S))| - |T(S)| = |B(S)| - |S| = |\partial S|.$$

Let  $\mathbf{x}$  be the element in  $S$  such that  $\mathbf{y} = T(\mathbf{x}, S)$ . We consider two cases.

*Case 1.*  $(x_i, x_j) \neq (1, 0)$ . In this case we certainly have  $\mathbf{y} = \mathbf{x}$ . Moreover, if  $\mathbf{z} \in \partial\mathbf{x}$  such that  $(z_i, z_j) = (1, 0)$  then it is readily checked that  $\mathbf{z} + \mathbf{e}_i + \mathbf{e}_j \in B(\mathbf{x})$ . Therefore the transformation  $T(\cdot, B(S))$  leaves  $B(\mathbf{x})$  invariant. Hence,  $B(\mathbf{x}) \subseteq T(B(S))$ .

*Case 2.*  $(x_i, x_j) = (1, 0)$ . Then  $\mathbf{z} = \mathbf{x} + \mathbf{e}_i + \mathbf{e}_j$  is the only neighbour of  $\mathbf{x}$  with  $(z_i, z_j) = (0, 1)$ .

*Case 2.1*  $\mathbf{y} = \mathbf{x}$ . Then, by the definition of  $T(\cdot, S)$ , we have  $\mathbf{z} \in S$  and  $T(\mathbf{z}, S) = \mathbf{z}$ . Observe that every neighbour  $\mathbf{z}'$  of  $\mathbf{x}$  is left invariant by  $T(\cdot, B(S))$ . This is clearly the case if  $(z'_i, z'_j) \neq (1, 0)$  and, if  $(z_i, z_j) = (1, 0)$ , because we then have  $\mathbf{z}'' = \mathbf{z}' + \mathbf{e}_i + \mathbf{e}_j \in B(\mathbf{z}) \subset B(S)$ . Hence

$$B(\mathbf{y}) = B(\mathbf{x}) = \cup_{\mathbf{z} \in B(\mathbf{x})} T(\mathbf{z}, B(S)) \subseteq T(B(S)).$$

*Case 2.2*  $\mathbf{y} \neq \mathbf{x}$ . Then  $\mathbf{y} \notin S$  but  $\mathbf{y} \in B(\mathbf{x}) \subseteq B(S)$ . Each neighbour  $\mathbf{z}$  of  $\mathbf{y}$  distinct from  $\mathbf{x}$  is of the form  $\mathbf{z} = \mathbf{z}' + \mathbf{e}_i + \mathbf{e}_j$  for some neighbour  $\mathbf{z}'$  of  $\mathbf{x}$  and therefore it belongs to  $T(B(S))$ . For  $\mathbf{x}$  itself we have  $T(\mathbf{x}, B(S)) = \mathbf{x}$  because  $\mathbf{y} = \mathbf{x} + \mathbf{e}_i + \mathbf{e}_j \in B(S)$ . Thus we again have  $B(\mathbf{y}) \subset T(B(S))$ . This completes the proof of (6).  $\square$

The weight of a vector  $\mathbf{x} \in \{0, 1\}^n$  is

$$w(\mathbf{x}) = \sum_{i=1}^n ix_i,$$

and the weight of a set  $S$  is

$$w(S) = \sum_{\mathbf{x} \in S} w(\mathbf{x}).$$

We note that, if  $i > j$  then  $w(T_{ij}(S)) \leq w(S)$ . Moreover, equality holds if and only if  $T_{ij}(S) = S$ . Thus, successive application of transformations  $T_{ij}$  using pairs  $i, j$  with  $i > j$  eventually produces a set which is stable by any of such transformations. This fact leads to the following definition.

**Definition 2.2.** *We say that a set  $S$  is compressed if  $T_{ij}(S) = S$  for each pair  $i, j \in [n]$  with  $i > j$ .*

Every set can be compressed by keeping its cardinality and without increasing its boundary. Therefore, in what follows we can restrict our attention to compressed sets in our study of optimal sets.

### 3. THE CASE $m = 2$

Theorem 1.2 follows from the following proposition which characterizes compressed optimal sets in  $J(n, 2)$ .

**Proposition 3.1.** *A set  $S$  of vertices of the graph  $J(n, 2)$  with cardinality  $\binom{t-1}{2} < |S| \leq \binom{t}{2}$  is optimal if and only if, up to isomorphism,  $S \subseteq \binom{[t]}{2}$ .*



*Proof.* Write  $V(J(n, 2))$  as the disjoint union

$$V(J(n, 2)) = S \cup \partial S \cup \tilde{S},$$

where  $\tilde{S}$  is the set of vertices at distance two from  $S$ . Let  $t = |\overline{S}|$  be the amount of elements of  $[n]$  in the support  $\overline{S}$  of  $S$ . The only vectors in  $\tilde{S}$  are the ones which have both nonzero coordinates in  $[n] \setminus \overline{S}$ . Therefore

$$|\tilde{S}| = \binom{n-t}{2},$$

and

$$|\partial S| = \binom{n}{2} - |S| - \binom{n-t}{2}.$$

Hence, for a given cardinality  $|S|$ ,  $|\partial S|$  is an increasing function of  $t$  alone. The optimal value is therefore obtained when  $t$  is smallest possible, which is the smallest  $t$  such that  $|S| \leq \binom{t}{2}$ . This is achieved for any subset  $S \subset \binom{[t]}{2}$  if  $|S| > \binom{t-1}{2}$ .  $\square$

As a consequence of the above proposition, we can see that the solution to the isoperimetric problem in  $J(n, 2)$  is unique (up to isomorphism) for sets of cardinality  $\binom{t}{2}$  (and also for sets of cardinality  $\binom{t}{2} - 1$ ).

#### 4. SMALL SETS

In this section we prove Theorem 1.3. Consider the partition

$$V(J(n, m)) = \{\mathbf{x} \in V(J(n, m)) | x_n = 0\} \cup \{\mathbf{x} \in V(J(n, m)) | x_n = 1\} = V_0 \cup V_1.$$

The subgraph of  $J(n, m)$  induced by  $V_0$  is isomorphic to  $J(n-1, m)$  and the subgraph induced by  $V_1$  is isomorphic to  $J(n-1, m-1)$ . There is an edge in  $J(n, m)$  joining  $\mathbf{x} \in V_0$  with  $\mathbf{y} \in V_1$  if and only if  $\overline{\mathbf{y}} \setminus \{n\} \subset \overline{\mathbf{x}}$ .

**Lemma 4.1.** *Let  $S$  be a set of vertices in  $J(n, m)$ . Let  $S_0 = S \cap \{x_n = 0\}$  and  $S_1 = S \cap \{x_n = 1\}$ . If  $S$  is compressed then*

$$B(S) = B(S_0) \text{ and } |B(S)| = |B'(S'_0)| + |\Delta(S'_0)|,$$

where  $S'_0 = \{\mathbf{x} \in \{0, 1\}^{n-1} : (\mathbf{x}, 0) \in S_0\}$  and  $B'$  denotes the ball in  $J(n-1, m)$ .

*Proof.* Let  $\mathbf{x} \in S_1$  and  $i \notin \overline{\mathbf{x}}$ . Since  $S$  is compressed, we have

$$\mathbf{y} = \mathbf{x} + \mathbf{e}_n + \mathbf{e}_i \in S_0,$$

which implies  $\mathbf{x} = \mathbf{y} + \mathbf{e}_n + \mathbf{e}_i \in B(S_0)$ . Hence,  $S_1 \subset B(S_0)$ . Moreover, if  $j \in \overline{\mathbf{x}}$  then

$$\mathbf{x} + \mathbf{e}_i + \mathbf{e}_j = (\mathbf{x} + \mathbf{e}_n + \mathbf{e}_i) + \mathbf{e}_n + \mathbf{e}_j \in B(S_0),$$

so that  $B(\mathbf{x}) \subset B(S_0)$ . Hence  $B(S) = B(S_0) \cup B(S_1) = B(S_0)$ . This proves the first part of the sentence.

For the second part, we just note that  $B(S_0)$  is the disjoint union  $(B(S_0) \cap V_0) \cup (B(S_0) \cap V_1)$ . Since the subgraph induced by  $V_0$  is isomorphic to  $J(n-1, m)$ , we

have  $|B(S_0) \cap V_0| = |B'(S'_0)|$ . On the other hand, there is an edge in  $J(n, m)$  joining  $x \in V_1$  with  $y \in V_0$  if and only if  $\bar{x} \setminus \{n\} \subset \bar{y}$ . It follows that  $B(S_0) \cap V_1 = \Delta(S'_0)$ .  $\square$

The next proposition considers the case of Johnson graphs  $J(n, m)$  when  $n = 2m - 2$ .

**Proposition 4.2.** *For each  $k = 1, \dots, \binom{2m-2}{m}$  the initial segment of length  $k$  in the colex order is an optimal set of the graph  $J(2m - 2, m)$ .*

*Proof.* Let us recall (1) and express the ball of a set  $S$  as

$$B(S) = \nabla(\Delta(S)).$$

By the Kruskal–Katona theorem,  $\Delta S$  is minimized by the initial segment  $I_k$ ,  $k = |S|$ , in the colex order of  $\binom{[2m-2]}{m}$ . Moreover,  $\Delta I_k$  is also an initial segment of the colex order in  $\binom{[2m-2]}{m-1}$ .

The automorphism of the Boolean lattice given by taking complements sends the middle level  $\binom{[2m-2]}{m-1}$  to itself, and the initial segments in the colex order are exchanged by the initial segments in the lexicographic order. It follows that the initial segments of the colex order are also a solution to the minimization of the *upper* shadow (as well as the initial segments in the lexicographic order). Hence,  $I_k$  minimizes  $|\nabla(\Delta(S))|$ .  $\square$

We use the Proposition 4.2 as the base case of the induction for the proof of Theorem 1.3.

*of Theorem 1.3.* The proof is by induction on  $n$ . For  $n = 2m - 2$  the result follows from Proposition 4.2. Let  $S$  be a compressed set of cardinality  $k \leq m - 1$  in  $J(n, m)$ ,  $n \geq 2m - 1$  and consider its decomposition  $S = S_0 \cup S_1$ . Since  $S$  is compressed, every element in  $S_1$  gives rise to at least  $m - 1$  elements in  $S$ . Since  $k < m - 1$  we have  $S_1 = \emptyset$ .

By Lemma 4.1, the cardinality of the ball of  $S$  is

$$|B(S)| = |B'(S'_0)| + |\Delta(S'_0)|.$$

By the induction hypothesis, the initial segment in the colex order minimizes the ball  $|B'(S'_0)|$  in  $J(n - 1, m)$  as well as, by the Kruskal–Katona theorem, the lower shadow  $|\Delta(S'_0)|$ .  $\square$

## 5. OPTIMAL SETS FOR LARGE $n$

In this Section we give the proof of Theorem 1.4. In what follows we call a positive integer  $k$  *critical* if its  $m$ -binomial representation has length  $m$  (namely, it has  $r = m - 1$ ).

of Theorem 1.4. Let  $S$  be an optimal set with cardinality  $k$  in  $J(n, m)$ . We may assume that  $S$  is compressed. Let  $n_0$  be such that the support of every element in  $S$  is contained in  $[n_0]$ . Since  $S$  is compressed, if the support of  $\mathbf{x} \in S$  contains  $n_0$  then we have  $T_{n_0 i}(\mathbf{x}) \in S$  for each  $i \in [n_0 - 1] \setminus \overline{\mathbf{x}}$ . It follows that  $n_0 \leq m + k + 1$ . For each  $n \geq n_0$  every element in  $\Delta(S)$  gives rise to  $n - n_0$  distinct vectors in  $\partial S$  which have a coordinate in  $[n_0 + 1, n]$  and therefore are disjoint from the ball  $B_0(S)$  in  $J(n_0, m)$ . Moreover every two such vectors which only differ in their coordinate from  $[n_0 + 1, n]$  come from a unique element in  $\Delta(S)$ . Therefore, we have

$$|B(S)| = |B_0(S)| + (n - n_0)|\Delta S|,$$

where  $B_0$  denotes the ball of  $S$  in  $J(n_0, m)$  and  $B$  denotes the ball of  $S$  in  $J(n, m)$ . Similarly, for  $I$  be the initial segment of length  $k$  in the colex order. We have

$$|B(I)| = |B_0(I)| + (n - n_0)|\Delta I|$$

Hence,

$$|B(S)| = |B(I)| + (|B_0(S)| - |B_0(I)|) + (n - n_0)(|\Delta(S)| - |\Delta(I)|). \quad (7)$$

If  $|\Delta(S)| > |\Delta(I)|$  then we have  $|B(S)| > |B(I)|$  for each sufficiently large  $n$ . By Theorem 1.6, if the  $m$ -binomial representation of  $k$  has less than  $m$  terms then the initial segment in the colex order is the unique solution to the Shadow Minimization Problem. It follows that, if  $S \neq I$  then  $|B(S)| > |B(I)|$  for all  $n > n_0 - |B_0(S)| + |B_0(I)|$ . This proves the first part of Theorem 1.4 and gives the estimate

$$n(k, m) \leq m + k + 1 - \mu_{m+k+1, m}(k) + f(k, m + k + 1, m).$$

On the other hand, we have  $\mu_{n, m}(k) \leq \mu_{n, m}(k + 1) - 1$ , since otherwise an optimal set  $X$  with cardinality  $k + 1$  satisfies  $|\partial(X \setminus \{x\})| < \mu_{n, m}(k)$  for every  $x \in X$ , a contradiction with the definition of  $\mu_{n, m}$ . Suppose that  $k$  is a critical integer and let  $k, k - 1, \dots, k - \ell + 1$  be the longest decreasing sequence of critical integers. By the above remark we have  $\mu_{n, m}(k) \leq \mu_{n, m}(k - \ell) + \ell$  and  $\mu_{n, m}(k - \ell) = f(k - \ell, n, m)$ , the cardinality of the boundary of an initial segment in colex order with length  $k - \ell$ . The value of  $\ell$  is clearly  $k_r$ . This proves the second part of the statement.  $\square$

## 6. INITIAL SEGMENTS WHICH ARE NOT OPTIMAL

We conclude the paper by proving Proposition 1.5, which shows that there are values of  $k$  for which the initial segment of length  $k$  in the colex order fails to be an optimal set of  $J(n, m)$  for all sufficiently large  $n$ .

We prove first that, for each  $m$  and each integer  $t$  sufficiently large with respect to  $m$ ,

$$g(t, m) = \binom{t}{m} + 3 \binom{t}{m-1}$$

is a critical cardinality, namely, the  $m$ -binomial expansion of  $g(t, m)$  has  $m$  terms. This means that the solution of the Minimal Shadow Problem is not unique for  $k = g(t, m)$ . This fact is used in the proof of Proposition 1.5.

**Lemma 6.1.** *There is an infinite strictly increasing integer sequence  $\{\lambda_i\}_{i \geq 0}$ ,  $\lambda_i + 1 < \lambda_{i+1}$  such that, for each  $t$  and each  $m \geq 1$ ,*

$$g(t, m) = \sum_{i=0}^{m-1} \binom{t - \lambda_i}{m - i} + 1. \quad (8)$$

*Proof.* By induction on  $m$ . For  $m = 1$  we have

$$g(t, 1) = t + 3 = \binom{t + 2}{1} + 1, \quad (9)$$

and for  $m = 2$ ,

$$g(t, 2) = \binom{t + 2}{2} + \binom{t - 2}{1} + 1, \quad (10)$$

giving  $\lambda_0 = -2$  and  $\lambda_1 = 2$ . By using  $\binom{n}{m} = \sum_{j=0}^{n-1} \binom{j}{m-1}$  and induction, for  $m \geq 3$  we have

$$\begin{aligned} g(t, m) &= \binom{t}{m} + 3 \binom{t}{m-1} \\ &= \sum_{j=0}^{t-1} \binom{j}{m-1} + 3 \sum_{j=0}^{t-1} \binom{j}{m-2} \\ &= \sum_{j=0}^{t-1} g(j, m-1) \\ &= \sum_{j=0}^{t-1} \left( \sum_{i=0}^{m-2} \binom{j - \lambda_i}{m-1-i} + 1 \right) \\ &= \sum_{i=0}^{m-2} \sum_{j=0}^{t-1} \binom{j - \lambda_i}{m-1-i} + t. \end{aligned}$$

By using  $\lambda_0 = -2$  and  $\lambda_i > 0$  for  $i \in [1, m-2]$ , we can write

$$\begin{aligned} g(t, m) &= \sum_{j=0}^{t-1} \binom{j - \lambda_0}{m-1} + \sum_{i=1}^{m-2} \left( \sum_{j=\lambda_i}^{t-1} \binom{j - \lambda_i}{m-1-i} + \sum_{j=0}^{\lambda_i-1} \binom{j - \lambda_i}{m-1-i} \right) + t \\ &= \binom{t - \lambda_0}{m} + \sum_{i=1}^{m-2} \binom{t - \lambda_i}{m-i} + t + \sum_{i=1}^{m-2} \sum_{j=0}^{\lambda_i-1} \binom{j - \lambda_i}{m-1-i}, \end{aligned}$$

which shows that (8) holds with

$$\begin{aligned}
\lambda_{m-1} &= - \sum_{i=1}^{m-2} \sum_{j=0}^{\lambda_i-1} \binom{j - \lambda_i}{m-1-i} + 1 \\
&= - \sum_{i=1}^{m-2} \sum_{\ell=1}^{\lambda_i} \binom{-\ell}{m-i-1} + 1 \\
&= - \sum_{i=1}^{m-2} \sum_{\ell=1}^{\lambda_i} (-1)^{m-i-1} \binom{m-i-2+\ell}{m-i-1} + 1 \\
&= \sum_{i=1}^{m-2} (-1)^{m-i} \binom{m-i-1+\lambda_i}{m-i} + 1 \\
&= \sum_{j=2}^{m-1} (-1)^j \binom{\lambda_{m-j} + j - 1}{j} + 1 \tag{11}
\end{aligned}$$

We observe that the sequence is uniquely determined once  $\lambda_1$  is fixed. The first values of the sequence are

$$-2, 2, 4, 7, 14, 51, 928, 409625, \dots$$

It remains to show that the sequence is increasing. We will in fact show that  $\lambda_m \geq \max\{\lambda_{m-1} + 2, \lambda_{m-1}^2/4\}$  for all  $m \geq 2$ . The above inequality holds for  $m \leq 7$  as shown by the first values of the sequence. By using (11) (with  $\lambda_m$  instead of  $\lambda_{m-1}$ ), we have,

$$\lambda_m \geq \sum_{j=2, j \text{ even}}^{2\lfloor m/2 \rfloor - 1} \left( \binom{\lambda_{m-j+1} + j - 1}{j} - \binom{\lambda_{m-j} + j}{j+1} \right).$$

For  $j = 2$  we have

$$\begin{aligned}
\binom{\lambda_{m-1} + 1}{2} - \binom{\lambda_{m-2} + 2}{3} &= \frac{\lambda_{m-1}^2}{4} + \frac{\lambda_{m-1}(\lambda_{m-1} + 2)}{4} - \binom{\lambda_{m-2} + 2}{3} \\
&\geq \frac{\lambda_{m-1}^2}{4} + \frac{\lambda_{m-2}^4}{64} + \frac{2\lambda_{m-2}^2}{16} - \binom{\lambda_{m-2} + 2}{3} \\
&> \frac{\lambda_{m-1}^2}{4}
\end{aligned}$$

where the last inequality holds (for  $m \geq 4$ ) because the largest root of the polynomial  $x^4/64 + x^2/8 - \binom{x+2}{3}$  is smaller than  $\lambda_4 = 14$ .

On the other hand, for  $j \geq 4$ ,

$$\begin{aligned} \binom{\lambda_{m-j+1} + j - 1}{j} - \binom{\lambda_{m-j} + j}{j+1} &= \frac{1}{j!} \left( \prod_{t=0}^{j-1} (\lambda_{m-j+1} + t) - \frac{\prod_{t=0}^j (\lambda_{m-j} + t)}{j+1} \right) \\ &\stackrel{\lambda_{m-j+1} > \lambda_{m-j} + 1}{>} \frac{1}{j!} \prod_{t=2}^j (\lambda_{m-j} + t) \left( \lambda_{m-j+1} - \frac{\lambda_{m-j}(\lambda_{m-j} + 1)}{j+1} \right) \\ &\stackrel{\lambda_{m-j+1} \geq \lambda_{m-j}^2/4}{\geq} \frac{1}{j!} \prod_{t=2}^j (\lambda_{m-j} + t) \left( \frac{\lambda_{m-j}^2}{4} - \frac{\lambda_{m-j}(\lambda_{m-j} + 1)}{j+1} \right). \end{aligned}$$

The right-hand side is nonnegative if  $m - j \geq 2$ , as then  $\lambda_{m-j} \geq 4$ . If  $m - j = 1$  then it follows by induction on  $j \geq 1$  that

$$\binom{\lambda_2 + j - 1}{j} - \binom{\lambda_1 + j}{j+1} = \binom{3+j}{j} - \binom{2+j}{j+1} > 0.$$

This completes the proof.  $\square$

For  $t$  larger than  $\lambda_{m-1}$  the equality (8) in Lemma 6.1 provides the  $m$ -binomial expansion of  $g(t, m)$ . Hence this binomial expansion has length  $m$  and, by Theorem 1.6,  $g(t, m)$  is a critical cardinality. The proof of Proposition 1.5 uses this fact by choosing two distinct optimal sets for the SMP problem which have different boundaries in the Johnson graph.

*Proof. of Proposition 1.5* Let  $S = B_0(\binom{[t]}{m})$ , the ball of  $\binom{[t]}{m}$  in the Johnson graph  $J(n_0, m)$ ,  $n_0 = t + 3$ . The cardinality of  $S$  is

$$k = \binom{t}{m} + 3 \binom{t}{m-1} = g(t, m).$$

Let  $I(k)$  denote the initial segment of length  $k$  in the colex order.

By Lemma 6.1 the  $m$ -binomial expansion of  $k$  has  $m$  terms and can be written as

$$\binom{t}{m} + 3 \binom{t}{m-1} = \binom{t - \lambda_0}{m} + \binom{t - \lambda_1}{m-1} + \dots + \binom{t - \lambda_{m-1}}{1}. \quad (12)$$

We note that the shadows of  $S$  and of  $I(k)$  have the same cardinality:

$$\begin{aligned} |\Delta S| &= \binom{t}{m-1} + 3 \binom{t}{m-2} \\ &= \binom{t - \lambda_0}{m-1} + \binom{t - \lambda_1}{m-2} + \dots + \binom{t - \lambda_{m-2}}{1} + \binom{t - \lambda_{m-1}}{0} \\ &= |\Delta(I(k))|. \end{aligned}$$

It can be readily checked that the boundary of  $S$  in  $J(n_0, m)$  has cardinality

$$|\partial S| = \left| \partial^2 \binom{[t]}{m} \right| = 3 \binom{t}{m-2},$$

On the other hand, the boundary in  $J(n_0, m)$  of the initial interval  $I(k)$  as given by the function  $f(k, n_0, m)$  in (3) is

$$\begin{aligned} |\partial I(k)| &= \binom{t - \lambda_0}{m - 1} + \sum_{i=1}^{m-1} \left( \binom{t - \lambda_i}{m - i - 1} (n - t + \lambda_0 - 1) - \binom{t - \lambda_i}{m - i} \right) \\ &= \binom{t + 2}{m - 1} - \sum_{i=1}^{m-1} \binom{a_i}{m - i} = \binom{t + 2}{m - 1} + \binom{t + 2}{m} - \binom{t}{m} - 3 \binom{t}{m - 1} \\ &= \binom{t + 1}{m - 2} + 2 \binom{t}{m - 2}, \end{aligned}$$

which is strictly larger than  $3 \binom{t}{m-2}$ . Thus  $|B_0(I(k))| > |B_0(S)|$ . Moreover, by (7), for all  $n \geq n_0$ , we have

$$|B(I(k))| = |B(S)| + (|B_0(I(k))| - |B_0(S)|) > |B(I(k))|.$$

Therefore the initial segment in colex order  $I$  fails to be an optimal set for all  $n \geq t + 3$ .  $\square$

#### ACKNOWLEDGEMENTS

The authors are grateful to the comments and remarks of the referees, pointing out some inaccuracies in the original manuscript and helping to improve the readability of the paper.

#### REFERENCES

- [1] A. C. A.E. Brouwer and A. Neumaier. *Distance-Regular Graphs*. Springer-Verlag, 1989.
- [2] R. Ahlswede and G. Katona. Graphs with maximal number of adjacent pairs of edges. *Acta Math. Acad. Sci. Hungar.*, 32:97–120, 1978.
- [3] C. Bey. Remarks on an edge-isoperimetric problem. In *General theory of information transfer and combinatorics, Lecture Notes in Comput. Sci.*, 4123, pages 971–978. Springer, Berlin, 2006.
- [4] S. L. Bezrukov. An isoperimetric problem for manhattan lattices. In *Proc. Int. Conf. Sets, Graphs and Numbers, Budapest*, pages 2–3, 1991.
- [5] S. L. Bezrukov and U. Leck. A simple proof of the Karakhanyan-Riordan theorem on the even discrete torus. *SIAM J. Discrete Math.*, 23(3):1416–1421, 2009.
- [6] S. L. Bezrukov and O. Serra. A local-global principle for vertex-isoperimetric problems. *Discrete Math.*, 257(2-3):285–309, 2002. Kleitman and combinatorics: a celebration (Cambridge, MA, 1999).
- [7] B. Bollobás. *Combinatorics*. Cambridge University Press, Cambridge, 1986. Set systems, hypergraphs, families of vectors and combinatorial probability.
- [8] B. Bollobás and I. Leader. Compressions and isoperimetric inequalities. *J. Combin. Theory Ser. A*, 56(1):47–62, 1991.
- [9] B. Bollobás and I. Leader. Isoperimetric problems for  $r$ -sets. *Combin. Probab. Comput.*, 13(2):277–279, 2004.
- [10] P. J. Cameron. *Permutation groups*, volume 45 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1999.
- [11] D. Christofides, D. Ellis, and P. Keevash. An approximate isoperimetric inequality for  $r$ -sets. *Electron. J. Combin.*, 20(4):Paper 15, 12, 2013.

- [12] P. Erdős, C. Ko, and R. Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 12:313–320, 1961.
- [13] P. Frankl and Z. Füredi. A short proof for a theorem of Harper about Hamming-spheres. *Discrete Math.*, 34(3):311–313, 1981.
- [14] Z. Füredi and J. R. Griggs. Families of finite sets with minimum shadows. *Combinatorica*, 6(4):355–363, 1986.
- [15] C. D. Godsil. *Algebraic combinatorics*. Chapman and Hall Mathematics Series. Chapman & Hall, New York, 1993.
- [16] C. D. Godsil. *Algebraic combinatorics*. Chapman and Hall Mathematics Series. Chapman & Hall, New York, 1993.
- [17] L. H. Harper. Optimal numberings and isoperimetric problems on graphs. *J. Combinatorial Theory*, 1:385–393, 1966.
- [18] L. H. Harper. *Global methods for combinatorial isoperimetric problems*, volume 90 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2004.
- [19] S. Hart. A note on the edges of the  $n$ -cube. *Disc. Math.*, 14:157–163, 1976.
- [20] V. M. Karachanjan. A discrete isoperimetric problem on multidimensional torus. *Doklady AN Arm. SSR*, 74(2):61–65, 1982.
- [21] G. Katona. A theorem of finite sets. In *Theory of graphs (Proc. Colloq., Tihany, 1966)*, pages 187–207. Academic Press, New York, 1968.
- [22] J. B. Kruskal. The number of simplices in a complex. In *Mathematical optimization techniques*, pages 251–278. Univ. of California Press, Berkeley, Calif., 1963.
- [23] I. Leader. Discrete isoperimetric inequalities. In *Probabilistic combinatorics and its applications (San Francisco, CA, 1991)*, volume 44 of *Proc. Sympos. Appl. Math.*, pages 57–80. Amer. Math. Soc., Providence, RI, 1991.
- [24] M. Mörs. A generalization of a theorem of Kruskal. *Graphs Combin.*, 1(2):167–183, 1985.
- [25] O. Riordan. An ordering on the even discrete torus. *SIAM J. Discrete Math.*, 11(1):110–127, 1998.
- [26] J. K. S.M. Cioabă and W. Li. Disconnecting strongly regular graphs. *European J. Combinatorics*, 38:1–11, 2014.
- [27] K. K. S.M. Cioabă and J. Koolen. On a conjecture of brouwer involving the connectivity of strongly regular graphs. *Journal of Combinatorial Theory, Series A*, 119:904–922, 2012.

DEPARTMENT OF MATHEMATICS, UNIVERSITAT POLITÈCNICA DE CATALUNYA, BARCELONA

*E-mail address:* victor.diego@upc.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITAT POLITÈCNICA DE CATALUNYA AND BARCELONA  
GRADUATE SCHOOL OF MATHEMATICS

*E-mail address:* oriol.serra@upc.edu

COMPUTER SCIENCE INSTITUTE OF CHARLES UNIVERSITY (IUUK AND ITI), PRAGUE

*E-mail address:* lluis.vena@gmail.com